

OEC040

Artificial Intelligence Generativa e Microsoft Azure AI

Durata: 2 gg

Descrizione

Il corso ha un taglio doppio: applicativo e tecnologico. Nella prima parte, dopo una breve panoramica sulle principali app e siti che utilizzano AI generativa a scopo ludico, dimostrativo o di curiosità, il corso si concentra sull'utilizzo della AI generativa nel mondo del lavoro e per le aziende, ed illustra molti casi d'uso aziendali che si stanno ormai consolidando nel mondo, soprattutto nei campi del finance, e-health, cyber-security, CRM & marketing, image editing, programming coding, cripto-valute NFT.

Nella seconda parte, più tecnologica appunto, il corso illustra, in modo generale ed intuitivo, le complesse, recentissime e super-innovative tecnologie di Deep Learning utilizzate dalla AI generativa: in primis le reti neurali ricorrenti (RNN), i trasformatori (transformer), le reti neurali avversarie (GAN) e le loro ultime evoluzioni. Sono forniti anche numerosi link per possibili successivi approfondimenti.

Il corso si pone inoltre l'obiettivo di familiarizzare i partecipanti con la terminologia del settore, spesso criptica e ridondante, aiutandolo ad orientarsi nell'ampio panorama di soluzioni e strumenti.

Il corso demistifica queste nuove tecnologie di AI generativa, cioè ne illustra in modo laico, al di là dell'hype a volte persino eccessivo, le opportunità e potenzialità, che appaiono obiettivamente amplissime. Dietro la AI generativa non c'è la "magia" nè tantomeno la "intelligenza" (questo termine è ormai ritenuto fuorviante), ma matematica e statistica avanzate (in questo corso solo accennate, come detto). Il corso illustra anche i limiti (attuali?) di queste tecnologie ed i rischi, sia di tipo giuridico che etico, che sono insiti in un loro uso improprio, già dietro l'angolo.

Tutti i casi d'uso aziendali illustrati sono accompagnati da codice software ben commentato e funzionante, che potrà costituire un buon punto di partenza per eventuali successive personalizzazioni da parte dei partecipanti. Il corso non approfondisce il contenuto tecnico del codice software (se non su richiesta), che è usato solo come demo concreta delle funzionalità di AI generativa. Il corso elenca ma non illustra i casi d'uso non-aziendali della AI generativa: arte, poesia, musica.

In sintesi: il corso fornisce le idee principali e gli strumenti per partire con la AI generativa (kick-off projects).

Il corso dedica uno spazio ampio ma non esclusivo alle soluzioni tecniche di AI generativa di Microsoft Azure AI (in particolare, i Cognitive Services)

Nel corso sono mostrate modalità alternative legali di accesso al servizio ChatGPT, che non risentono del blocco in Italia dal 30 marzo 2023.

Vedi anche: ChatGPT - Crash Course.

Vedi anche: corso di approfondimento su Azure Open AI e ChatGPT: scopi ed utilizzo.

L'intelligenza artificiale: di cosa si tratta

Una nuova onda, da fine 2022, sta percorrendo i media, la società e le aziende, anche in Italia: l'intelligenza artificiale generativa, che permette di produrre, da un input testuale, vari tipi di contenuti multimediali: immagini, video, audio, altro testo, grafici, codice software e 3D. Cioè l'intelligenza artificiale - o meglio il machine learning

e deep learning che ne sono il suo cuore - non si limita più a prevedere, classificare o spiegare i dati, ma ne genera appunto di nuovi. Ed ancora più sorprendente e pieno di conseguenze è che questi nuovi dati possano essere espressi in qualsiasi lingua umana, ed anche in linguaggi come il codice software, la musica e le sequenze di proteine.

Anche gli osservatori più prudenti ammettono ormai che la AI generativa ha tutte le caratteristiche per risultare il nuovo "evento disrupting", in grado cioè di cambiare profondamente il modo di lavorare dei prossimi decenni, come lo sono stati in passato i PC ed Internet. Secondo IBM ed Harvard a breve l'AI generativa costituirà una delle competenze chiave in molti curricula professionali. Secondo Microsoft l'AI generativa è "la prossima frontiera".

Il prodotto che ha fatto da catalizzatore di questa nuova onda è stato senz'altro ChatGPT di OpenAI (in questo corso brevemente illustrato, per il quale si rimanda al corso OEC039 per maggiori approfondimenti), ma in realtà la pleora di strumenti di AI generativa disponibili è oggi molto più ampia: da Dall-e di OpenAI stessa a BERT e Bard di Google a BART di Facebook, e molti altri ancora. ChatGPT infatti si focalizza soprattutto sulla generazione di testi da testi (sequence-to-sequence model), altri strumenti risultano più efficaci nella generazione di immagini e video da testo, ad esempio, oppure nella trasformazione di immagini e video in testo. Il corso presenta questi e molti altri strumenti di AI generativa, li cala nei loro possibili scenari di concreto utilizzo in azienda e li confronta nelle loro caratteristiche e performance.

A chi è rivolto?

Business Manager, Project Manager, Data Scientist, Analisti, Sviluppatori. In generale a chiunque si voglia fare un'idea a 360° di questa nuova tecnologia e delle sue applicazioni aziendali pratiche.

Prerequisiti

Non è necessario alcun prerequisito.

Contenuti

Parte 1: Introduzione

Cos'è l'AI generativa e le sue tre aree principali:

- generazione di contenuti
- trasformazione di contenuti
- arricchimento dei dataset

La storia dell'AI generativa

I tre driver storici dell'AI generativa: grandi dataset multimediali, nuovi algoritmi, maggiore potenza computazionale anche in periferia (l'edge)

Il panorama attuale della AI generativa (anche in Italia)

Cos'è l'AI conversazionale

I tipi di dati generabili: testo, audio, video, immagini, 3D, strutturati (tabellari), codice software, serie temporali, musica.

Le trasformazioni Text2X, X2Text

Cosa sono i dati "sintetici": il game-changer in campo sanitario e nella prevenzione delle frodi

Casi d'uso aziendali di AI generativa: prima rassegna

L'AI generativa come prossimo sostituto dei motori di ricerca? (Google Search integrato con Bard vs Microsoft

Bing integrato con OpenAI)?

Parte 2: il Natural Language Processing (NLP) ed i Large Language Model (LLM)

Cos'è il Natural Language Processing ed i suoi casi d'uso principali:

- comprensione di un testo
- sintesi di un testo
- traduzione di lingua
- generazione di dialoghi
- generazione di risposte a domande
- completamento automatico di testi parziali
- riconoscimento e sintesi vocale
- riconoscimento di cifre scritte a mano
- l'analisi del sentimento
- image captioning
- tagging di parti del discorso (POS tagging)
- OCR (Optical Character Recognition)
- riconoscimento vocale
- named entity recognition: il livello semantico

I metodi deterministici e probabilistici

Modelli di linguaggio classici: a borsa di parole (bag-of-words), n-grammi, modello di Markov, regressione logistica

Cos'è un modello linguistico di grandi dimensioni (Large Language Model - LLM)

I Neural Language Model: il cuore della AI generativa

- Word2vec
- Word embedding pre-addestrati
- Reti Neurali Ricorrenti (RNN)
- LSTM
- Meccanismo di attenzione
- Meccanismo di auto-attenzione e di attenzione multi-head
- Encoder / Decoder
- Trasformatori: pre-allenamento e fine-tuning

Parte 3: la Computer Vision generativa

Cos'è la Computer Vision:

- collegare immagini e parole (tagging e captioning)
- generare nuove immagini e nuovi volti
- trasformare le immagini (domain transfer e style transfer)
- generare nuovi dati sintetici con le Generative Adversarial Network (GAN): le parti mancanti di un'immagine, immagini di più alta risoluzione grafica, immagini trasformate.
- classificare immagini con le CNN (Convolutional Neural Network)
- rilevare gli oggetti in un'immagine (object detection)
- a cosa serve la Data Augmentation

Le GAN più recenti: DCGAN e GAN di Wasserstein

Parte 4: gli strumenti di AI generativa

Rassegna dei principali strumenti di AI generativa sul mercato:

- GPT di OpenAI per la generazione di testi.
- BERT: la framework pre-addestrata di Google, utile per una vasta gamma di task di AI testuale, in particolare per compiti di classificazione, sentiment analysis, named entity recognition, previsione della frase successiva, etichettatura dei ruoli semantici, classificazione del testo, coreference resolution. Oggi è usata da Google Search.
- BART: la framework di Facebook (una generalizzazione di GPT e BERT)
- LaMDA di Google
- DALL-E di OpenAI
- ChatGPT di OpenAI
- Bard di Google: il principale rivale di ChatGPT?
- Wenxin, la risposta cinese a ChatGPT
- Synthesia, per generare video da un testo scritto dall'utente
- Amazon Polly, il software per conversione da testo a voce degli Amazon Web Services
- Midjourney e Adobe Stable Diffusion per la generazione di immagini
- Play.ht, per la generazione di audio e voci
- ChatBCG, per la creazione di presentazioni PowerPoint
- Looka e PFP maker per le startup
- LensAI, l'app virale per i ritratti personali
- La soluzione IBM di AI generativa

La tecnologia open-source per l'AI generativa:

- Python / Jupyter / PyTorch: esame di codice
- TensorFlow / Keras / R / RStudio

ChatGPT fai-da-te (in Python con una singola GPU)

Hardware per AI generativa: GPU, TPU, Google Colab

Parte 5: la generazione / correzione del codice software con l'AI generativa

Rassegna dei principali strumenti:

- GitHub Copilot e GitHub CodeAssist (GitHub è una società di Microsoft)
- Codex di OpenAI
- Durable

Parte 6: limiti e rischi della AI generativa

- un esame dettagliato
- bias nei dati di allenamento
- deepfake di Striscia la notizia)
- etica nella AI generativa

Parte 7: focus sulla tecnologia di AI generativa di Microsoft

- Microsoft Azure AI (focus sui servizi cognitivi generativi):
oper la computer vision
oper il parlato (speech)
oper il linguaggio (NLP)
- Microsoft e OpenAI
- ChatGPT integrato in Microsoft Bing vs ChatGPT nativo
- AI generativa in Microsoft Viva Sales
- Vall-E, il simulatore di voce di Microsoft